微分ゲームにおける
シュタッケルベルク均衡II

Stackelberg Equilibria
in a Differential Game II

藤　原　憲　二

This paper discusses a *global* Stackelberg equilibrium of the model developed in Fujiwara (2012). We provide a detailed derivation of this equilibrium.

Kenji Fujiwara

JEL：C73

キーワード：微分ゲーム、大域的シュタッケルベルク均衡

Keywords：dynamic game, global Stackelberg equilibrium

1　導入

本稿は藤原 (2012) の続編である。藤原 (2012) では生産補助金を決める政府を先導者、生産量を決める独占企業を追随者とする微分ゲーム・モデルを考え、そのモデルにおけるオープンループ・シュタッケルベルク均衡を導出した。またオープンループ・シュタッケルベルク均衡には特殊な場合を除いて時間不整合性の問題が起こることを指摘した。それを受けて本稿と次稿では全てのプレーヤーが初期時点に以後の戦略を予め決めてしまうオープンループ解ではなく、各時点でその時々の状態変数の値を見た上で戦略を決めるというフィードバック解を考えその解法を紹介する。

ただフィードバック解の導出はオープンループ解よりも込み入った計算が必要となる上に、フィードバック・シュタッケルベルク均衡もさらに 2 つに

細分化されるため、本稿ではそのうちの一方のみを取り上げ他方は次稿に譲る。その 2 つのフィードバック・シュタッケルベルク均衡は Basar and Olsder (1995)、Mehlmann (1988), Long (2010) にならい大域的シュタッケルベルク均衡 (global Stackelberg equilibrium) と段階的シュタッケルベルク均衡 (stagewise Stackelberg equilibrium) と呼ばれる。技術的な詳細は次節以降で述べるが、前者は追随者は毎時点に戦略を決め先導者は初期時点に戦略を決める（プレコミットする）のに対し、後者は両プレーヤーが毎時点に戦略を決める。したがって、前者では事実上先導者の行動はオープンループ解と同じになっている。このような解を「フィードバック」解と呼ぶことに違和感を持つ人もいるだろうが、それがなぜ「フィードバック」解と呼ばれるのかも次節以降で説明する。

本稿の構成は次の通りである。第 2 節では前稿で定式化したモデルを簡単に再掲する。第 3 節では問題を後方帰納法で解くために追随者の行動を記述する。第 4 節では先導者の行動を考え、大域的シュタッケルベルク均衡を導出する。第 5 節は結論である。

2 モデル

モデルは次の通りである。独占企業が自社の利潤の無限期間における割引現在価値を最大化するように生産量の時間流列 $x(t)$ を決める。他方、政府は独占（過小生産）による死荷重を是正するために、経済厚生の割引現在価値を最大化するように生産補助金率の時間流列 $s(t)$ を決める[1)]。ここで政府が先に s を決め、その後に企業が x を決めるというシュタッケルベルク・ゲームを考える。

このモデルでは財価格は瞬時に需要と供給が一致するようには決まらず、次の微分方程式で表されるように超過需要または超過供給に応じて緩慢に変化するとする。

$$\dot{p} = k(a - p - x), \quad k > 0. \tag{1}$$

1) 以下では特に混乱が生じない限り、時間を表す t は省略する。

ここで p は市場価格、$a-p$ は需要関数であり $a-p-x$ は超過需要を表す。藤原 (2012) での計算を行うと、各プレーヤーの目的関数は次のように表される。

$$\text{企業（追随者）}: \int_0^\infty e^{-rt}\left(px - cx + sx - \frac{x^2}{2}\right) dt \tag{2}$$

$$\text{政府（先導者）}: \int_0^\infty e^{-rt}\left[\frac{(a-p)^2}{2} + px - cx - \frac{x^2}{2}\right] dt. \tag{3}$$

(2) 式右辺の $(-cx - x^2/2)$ は生産費用を表し、(3) 式右辺の被積分関数の第 1 項は消費者余剰、第 2 項以降は生産補助金を除いた独占企業の利潤を表している。企業は (2) を最大化するように x を、政府は (3) を最大化するように s を決める。

3　追随者の行動

オープンループ解では両プレーヤー共に初期時点に以後の行動を全て決めてしまうことから、最大値原理を使って問題を解く。それに対してフィードバック解では（少なくとも一方のプレーヤーは）毎期問題を解き直すことを仮定するから、動学計画法によって問題を解くことになる。本稿で考える大域的シュタッケルベルク均衡の最大の特徴は先導者が $s(p) = \alpha p + \beta$ という状態変数 p の関数としての戦略を決めたとして追随者の問題を解く点にある。ここで 3 点注意しておく。第 1 に先導者の戦略を $\alpha p + \beta$ と p に関する 1 次関数にしているのは単に計算をやりやすくするためだけの仮定である。第 2 に先導者は追随者の行動を知った上で α, β という時間とは関係のない 2 つの定数を自分の初期時点から無限先までの目的関数の積分値（割引現在価値）を最大化するように決めるが、それは初期時点に決められるために先導者の問題は事実上オープンループ的なものとなっている。第 3 に、しかしながら得られた解を「フィードバック」解と呼ぶのは、先導者の戦略が $\alpha p + \beta$ という状態変数の関数になっているからである。

今、政府の戦略が状態変数の 1 次関数として $\alpha p + \beta$ であるとする。追随者である企業はこれを所与として問題を解く。そこで (2) の s の部分に $\alpha p + \beta$ を代入し、企業のハミルトン＝ヤコビ＝ベルマン方程式をたてると次のように

なる[2)]。

$$rV(p)=\max_x\left\{px-cx+(\alpha p+\beta)x-\frac{x^2}{2}+kV'(p)(a-p-x)\right\}. \tag{4}$$

ここで $V(p)$ は企業の価値関数である。右辺を x で微分してゼロとおくと $p-c+\alpha p+\beta-x-kV'(p)=0$ という 1 階条件を得る。ここで計算をやりやすくするために価値関数が $V(p)=Ap^2/2+Bp+C$ という 2 次関数であると仮定する。すると $V'(p)=Ap+B$ となりこの結果を先の 1 階条件に代入し、x について解くと次のようになる。

$$x=(-kA+\alpha+1)p-kB+\beta-c. \tag{5}$$

この結果を (4) の右辺に代入し整理すると、次のような p に関する恒等式を得る。

$$r\left(\frac{A}{2}p^2+Bp+C\right)\equiv\frac{[(-kA+\alpha+1)p-kB+\beta-c]^2}{2}+k(Ap+B)(-p+a).$$

これは恒等式であるから両辺の p^2, p にかかる係数と定数項は常に等しくなっていなければならない。そこで両辺の p^2, p にかかる係数と定数項を等しくおいた式を並べると次のようになる。

$$\begin{aligned}\frac{rA}{2}&=\frac{(-kA+\alpha+1)^2}{2}-kA\\ rB&=(-kA+\alpha+1)(-kB+\beta-c)+k(Aa-B)\\ rC&=\frac{(-kB+\beta-c)^2}{2}+kBa.\end{aligned}$$

定数 A,B,C はこの 3 本の連立方程式を解いて求めればよいが、最初の式だけで A が決まり、それを 2 番目の式に代入すると B が、さらにそれらを最後の式に代入すれば C が求められる[3)]。ここで最初の式は A に関する 2 次方程式なので、解の公式を用いて解を求めると次のように 2 つの実数解が得られる。

$$A=\frac{r+2k(\alpha+2)\pm\sqrt{\Delta}}{2k^2},\quad \Delta\equiv(r+2k)[r+2k(2\alpha+3)]>0.$$

ここで (5) を (1) に代入すると

---

2) 連続時間の動学計画法については Leonard and Long (1992), Dockner et al. (2000) を参照せよ。

3) C は後の議論で重要な役割を果たさないので本稿では明示的に求めない。

$$\dot{p} = k[a - (-kA + \alpha + 2)p],$$

となるからこの微分方程式の定常状態が安定的であるためには、p にかかる係数が負でなければならない。この安定性を満たすためには 2 つある A の解のうち

$$A = \frac{r + 2k(\alpha + 2) - \sqrt{\Delta}}{2k^2}, \tag{6}$$

しか採用することができない。この結果を使い B を求めると次のようになる。

$$B = \frac{\left[r + 2k(\alpha + 2) - \sqrt{\Delta}\right] a + \left(-r - 2k + \sqrt{\Delta}\right)(\beta - c)}{k\left(r + \sqrt{\Delta}\right)}. \tag{7}$$

このようにして求められた A, B を (5) に代入すると最終的に追随者の戦略を次のように求めることができる。

$$\begin{aligned} x(p) &= \alpha^* p + \beta^* \\ \alpha^* &= \frac{-r - 2k + \sqrt{\Delta}}{2k} \\ \beta^* &= \frac{-\left[r + 2k(\alpha + 2) - \sqrt{\Delta}\right] a + 2(r + k)(\beta - c)}{r + \sqrt{\Delta}}. \end{aligned} \tag{8}$$

この中で α^* は α、β^* は α, β という先導者の戦略に依存していることに注意せよ。つまり教科書的なシュタッケルベルク・ゲームと同じように、追随者の戦略は先導者の戦略に依存している。

4 先導者の行動と均衡

前節で追随者の行動が記述できたので、本節では先導者である政府の行動からシュタッケルベルク均衡を導出する。政府は初期時点から無限先までの経済厚生の割引現在価値を最大化するように α, β を決める。したがって最適な α, β を求めるに先立って政府が最大化する目的関数を求めなければならない。この作業は重要だが代表的な微分ゲームの文献である Dockner et al. (2000), Long (2010) でも詳しい計算過程は書かれていないので、以下では先導者の目的関数の求め方を詳説する。

追随者の戦略が (8) となることが分かっているから、それを (1) に代入す

ると次のようになる。

$$\dot{p} = k[a - p - \alpha^* p - \beta^*] = k\left[a - \beta^* - (\alpha^* + 1)\, p\right].$$

この微分方程式は定数係数 1 階線形微分方程式なので、次のように明示的に解を求めることができる。

$$p(t) = e^{-k(\alpha^*+1)t}\,(p_0 - \overline{p}) + \overline{p}. \tag{9}$$

ここで p_0 は歴史的に与えられた価格の初期値であり、$\overline{p}$ は価格の定常状態値であり次式で与えられる。

$$\overline{p} = \frac{a - \beta^*}{\alpha^* + 1} = \frac{[r + k(\alpha + 2)]\, a - (r + k)(\beta - c)}{r(\alpha + 2) + k(2\alpha + 3)}. \tag{10}$$

前稿または第 2 節でみたように政府の瞬時的目的関数は消費者余剰と生産補助金を除いた企業の利潤の和

$$W \equiv \frac{(a - p)^2}{2} + px - cx - \frac{x^2}{2},$$

であるが、この式の x に追随者の戦略 (8) を代入し式を展開すると次のようになる。

$$\begin{aligned} W = &\left(-\alpha^{*2} + 2\alpha^* + 1\right) p^2 + (-2\alpha^*\beta^* - 2c\alpha^* + 2\beta^* - 2a)\, p \\ &-\beta^{*2} - 2c\beta^* + a^2. \end{aligned}$$

この式の p に (9) を代入しさらに整理すると次のようになる。

$$\begin{aligned} \frac{W}{2} &= De^{\frac{r-\sqrt{\Delta}}{k}t}\,(p_0 - \overline{p})^2 + Ee^{\frac{r-\sqrt{\Delta}}{2k}t}\,(p_0 - \overline{p}) + F \\ D &\equiv -\alpha^{*2} + 2\alpha^* + 1 \\ E &\equiv \overline{p}\left(-\alpha^{*2} + 2\alpha^* + 1\right) - \alpha^*\beta^* - c\alpha^* + \beta^* - a \\ F &\equiv \overline{p}^2\left(-\alpha^{*2} + 2\alpha^* + 1\right) + 2\overline{p}(-\alpha^*\beta^* - c\alpha^* + \beta^* - a) - \beta^{*2} - 2c\beta^* + a^2. \end{aligned}$$

これらの式の $\overline{p}$ は (10) で与えられ α^*, β^* は (8) を通して α, β に依存しているから、D, E, F は全て先導者が決める α, β に依存していることが分かる。

このようにして求められた先導者の瞬時的目的関数に e^{-rt} をかけて 0 から ∞ までの定積分を計算すると先導者の目的関数が α, β の関数として求められる。その関数は t を含まず α, β, p_0 の関数となる。したがって先導者の目的関

数を最大にする α, β は得られた目的関数を α, β で偏微分してゼロとおくことで得られる。しかしそのようにして得られた 2 本の 1 階条件は極めて複雑な形をしており、明示的に解を求めることができない。それに加えて深刻なのはそのような 2 本の 1 階条件を満たす α, β はともに p_0 の関数となってしまう。これは、もし政府がある時点 $t > 0$ で問題を解き直すと、そのときに得られる α, β は時点 t の状態変数 $p(t)$ の実現値に依存してしまい、一般に $p_0 \neq p(t)$ であることからゼロ時点で解いた α, β と時点 t で解いた α, β が違う値をとることを意味する。これは前稿で説明した時間不整合性の問題である。

Fujiwara and Long (2011, 2012) ではこの問題を回避するために次のような条件をおき「時間整合性条件」と呼んだ。

$$\overline{p} = \frac{a - \beta^*}{\alpha^* + 1} = \frac{a + c}{2}. \tag{11}$$

この式の意味は次の通りである。左辺は今のモデルにおける価格（状態変数）の定常状態値であり、右辺の $(a + c)/2$ は社会的最適な価格（経済厚生を最大化する価格）である。したがって (11) は定常状態における市場価格が社会的最適な価格に等しいことを意味する。一般には市場均衡価格が社会的最適価格に等しくなる保証はないが、前稿で求めたオープンループ・シュタッケルベルク均衡における定常状態価格が $(a + c)/2$ になっていることを思い起こすとそれほど不自然な条件ではないであろう。

(11) の時間整合性条件を β^* について解くと

$$\beta^* = a - \frac{(\alpha^* + 1)(a + c)}{2},$$

となりこれを上で求めた D, E, F に代入して β^* を消去し式を整理すると先導者の瞬時的目的関数は次のように α^* （α^* は α に依存していることに注意せよ）だけの関数になる。

$$2W = \left(-\alpha^{*2} + 2\alpha^* + 1\right) e^{-2(\alpha^*+1)t} \left(p_0 - \frac{a + c}{2}\right)^2 + \frac{(a + c)^2}{2}.$$

これに e^{-rt} をかけて 0 から ∞ までの定積分を計算すると最終的に先導者の目的関数は次のように得られる。

$$2\int_0^\infty e^{-rt} W dt = \frac{-\alpha^{*2} + 2\alpha^* + 1}{r + 2(\alpha^* + 1)} \left(p_0 - \frac{a + c}{2}\right)^2 + \frac{(a - c)^2}{2r}.$$

この式の第 1 項の括弧の中も第 2 項も α に依存しないから、最終的に先導者は次の式を最大化するように α だけを決め、もうひとつの選択変数である β は時間整合性条件から得られる。

$$\frac{-\alpha^{*2}+2\alpha^{*}+1}{r+2(\alpha*+1)},\quad \text{ただし}\alpha^{*}=\frac{-r-2k+\sqrt{(r+2k)[r+2k(2\alpha+3)]}}{2k}.$$

複雑な形をしているが形式的にはこれは 1 変数関数なので、α で微分してゼロとおけば最適な α が求められ、それらを逐次代入すれば全てのプレーヤーの均衡における戦略が求められる。制約条件にある α^{*} を目的関数に代入して α だけの式に直すと次のようになる。

$$F(\alpha)\equiv\frac{-(r+2k)[r+2k(\alpha+3)]+(r+4k)\sqrt{(r+2k)[r+2k(2\alpha+3)]}+2k^{2}}{2k\left\{r(k-1)+\sqrt{(r+2k)[r+2k(2\alpha+3)]}\right\}}.$$

これを α で微分してゼロとおくと次のような α に関する最大化の 1 階条件が得られる。

$$\begin{aligned}0=F'(\alpha)&\frac{r+2k}{\sqrt{(r+2k)[r+2k(2\alpha+3)]}\left\{r(k-1)+\sqrt{(r+2k)\,[r+2k(2\alpha+3)]}\right\}^{2}}\\&\times\left\{r(k-1)\left(r+4k-\sqrt{(r+2k)[r+2k(2\alpha+3)]}\right)-2k[(r+2k)\alpha+k]\right\}.\end{aligned}$$

シュタッケルベルク均衡は上の式の最後にある大括弧の中がゼロになる α として決まるが、それは煩雑な式になっているため以下では $k=1$ の場合に焦点を絞ろう。このとき最適な α は次のように非常に簡単な形として求められる。

$$\alpha=-\frac{1}{r+2}. \tag{12}$$

これが現在のモデルにおける大域的シュタッケルベルク均衡である。これを政府の目的関数に代入すると、均衡における政府（先導者）の利得が次のように求められる。

$$\int_{0}^{\infty}e^{-rt}W\,dt=\frac{r+4-\sqrt{r^{2}+8r+8}}{4}\left(p_{0}-\frac{a+c}{2}\right)^{2}+\frac{(a-c)^{2}}{4r}. \tag{13}$$

次稿で示すがもうひとつのシュタッケルベルク均衡である段階的シュタッケルベルク均衡においても先導者の利得が (13) として求められることが分かる。これは我々が特殊な関数形を考えていること、および $k=1$ という制約的な

条件を課しているからだと思われる[4)]。

5 結論

本稿では 2 つのフィードバック・シュタッケルベルク均衡のうち、先導者が $\alpha p + \beta$ という戦略のうち α, β という 2 つの（時間に依存しない）変数を自らの目的関数の割引現在価値を最大化するように決めるという大域的シュタッケルベルク均衡の導出法を詳説してきた。ただしこの解概念には次のような問題点がある。

第 1 に解の導出が極めて煩雑になる上に、明示的に最適な α を求めることができない。我々の経験によればこの問題は、制約となっている微分方程式が 1 次式であり目的関数が 2 次関数となっている線形 2 次（linear-quadratic）モデルでは必ず発生する。線形 2 次モデルは公共経済学、環境経済学、資源経済学など非常に多くの分野で応用されていることを考えると、解を明示的に求められる技術的な手法が開発されることが求められる[5)]。

第 2 に先導者の問題が事実上オープンループとなっているため、オープンループ・シュタッケルベルク均衡と同じ時間不整合性問題が避けられない。Fujiwara and Long (2011, 2012) および本稿ではこの問題を回避するために「時間整合性条件」を課した。確かにこれを課すと時間不整合性は解消されるが、その条件を正当化する数学的・経済学的理由がない[6)]。要約すると大域的シュタッケルベルク均衡が明示的に求められ、かつ人為的な時間整合性条件を課す必要がないようにするには関数形を厳しく制限しなければならない。応用上最も多く使われている線形 2 次モデルで上の 2 つの問題点を解消することがこの分野における大きな課題である。

4) 実際、Fujiwara and Long (2011, 2012) では 2 つのフィードバック・シュタッケルベルク均衡でプレーヤーの利得が違うことを数値例によって示している。解析的にこのことが成立するかは未解明の問題であり、今後の検討課題である。

5) Fujiwara (2012) は目的関数を 1 次同次関数にして解が明示的に解けるひとつの例を与えている。

6) Fujiwara (2012) のモデルではこの条件なしで大域的シュタッケルベルク均衡を求めている。

参考文献

[1] Basar, T. and G. Olsder (1995), *Dynamic Noncooperative Game Theory,* San Diego: Academic Press.

[2] Dockner, E. J., S. Jorgensen, N. V. Long and G. Sorger (2000), *Differential Games in Economics and Management Science,* Cambridge: Cambridge University Press.

[3] Fujiwara, K. and N. V. Long (2011), 'Welfare Implications of Leadership in a Resource Market under Bilateral Monopoly,' *Dynamic Games and Applications*, 1, 479-497.

[4] Fujiwara, K. and N. V. Long (2012), 'Optimal Tariffs On Exhaustible Resources: The Case Of Quantity-Setting,' *International Game Theory Review*, 14, 1240004-1-1.

[5] Leonard, D. and N. V. Long (1992), *Optimal Control Theory and Static Optimization in Economics,* Cambridge: Cambridge University Press.

[6] Long, N. V. (2010), *A Survey of Dynamic Games in Economics* Singapore: World Scientific Publishing Co.

[7] Mehlmann, A. (1988), *Applied Differential Games*, New York: Plenum Press.

[8] 藤原憲二 (2012), '微分ゲームにおけるシュタッケルベルク均衡 I,' 経済学論究, 6, 163-172.